

03CG



OIKE

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/10/036,729

TIME: 12:02:38

Input Set : N:\Crf3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Middeldorp, Jaap Michiel.

7 (ii) TITLE OF INVENTION: Peptides and nucleic acid sequences  
8 related to the Epstein-Barr virus.

10 (iii) NUMBER OF SEQUENCES: 22

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Akzo-Nobel Patent Department

14 (B) STREET: 1300 Piccard Drive, Suite 206

15 (C) CITY: Rockville

16 (D) STATE: Maryland

17 (E) COUNTRY: USA

18 (F) ZIP: 20850

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: Patentin Release #1.0, Version #1.25

26 (vi) CURRENT APPLICATION DATA:

C--&gt; 27 (A) APPLICATION NUMBER: US/10/036,729

C--&gt; 28 (B) FILING DATE: 21-Dec-2001

30 (vii) PRIOR APPLICATION DATA:

31 (A) APPLICATION NUMBER: 08/415,838

32 (B) FILING DATE:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Gormley, Mary E.

36 (B) REGISTRATION NUMBER: 34,409

38 (2) INFORMATION FOR SEQ ID NO: 1:

40 (i) SEQUENCE CHARACTERISTICS:

41 (A) LENGTH: 538 base pairs

42 (B) TYPE: nucleic acid

43 (C) STRANDEDNESS: double

44 (D) TOPOLOGY: unknown

46 (ii) MOLECULE TYPE: DNA (genomic)

48 (vi) ORIGINAL SOURCE:

49 (A) ORGANISM: Epstein-Barr virus

52 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

54	CATGATGGCA	CGCCGGCTGC	CCAAGCCCAC	CCTCCAGGGG	AGGCTGGAGG	CGGATTTTCC	60
56	AGACAGTCCC	CTGCTTCCTA	AATTTCAGA	GCTGAACCAG	AATAATCTCC	CCAATGATGT	120
58	TTTTCGGGAG	GCTCAAAGAA	GTTACCTGGT	ATTTCTGACA	TCCCAGTTCT	GCTACGAAGA	180
60	GTACGTGCAG	AGGACTTTTG	GGGTGCCCTG	GCGCCAACGC	GCCATAGACA	AGAGGCAGAG	240
62	AGCCAGTGTG	GCTGGGGCTG	GTGCTCATGC	ACACCTTGGC	GGGTCATCCG	CCACCCCCGT	300
64	CCAGCAGGCT	CAGGCCGCCG	CATCCGCTGG	GACCGGGGCC	TTGGCATCAT	CAGCGCCGTC	360

ENTERED

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/10/036,729

TIME: 12:02:38

Input Set : N:\Crf3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

66 CACGGCCGTA GCCCAGTCCG CGACCCCTC TGTTTCTTCA TCTATTAGCA GCCTCCGGGC 420  
 68 CGCGACTTCG GGGGCGACTG CCGCCGCTC CGCCGCCGCA GCCGTCGATA CCGGGTCAGG 480  
 70 TGGCGGGGGA CAACCCACG ACACCGCCCC ACGCGGGGCA CGTAAGAAAC AGTAGCCC 538

## 72 (2) INFORMATION FOR SEQ ID NO: 2:

## 74 (i) SEQUENCE CHARACTERISTICS:

75 (A) LENGTH: 176 amino acids

76 (B) TYPE: amino acid

77 (C) STRANDEDNESS: single

78 (D) TOPOLOGY: linear

80 (ii) MOLECULE TYPE: peptide

## 82 (vi) ORIGINAL SOURCE:

83 (A) ORGANISM: Epstein-Barr virus

## 86 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

88 Met Ala Arg Arg Leu Pro Lys Pro Thr Leu Gln Gly Arg Leu Glu Ala  
 89 1 5 10 15  
 91 Asp Phe Pro Asp Ser Pro Leu Leu Pro Lys Phe Gln Glu Leu Asn Gln  
 92 20 25 30  
 94 Asn Asn Leu Pro Asn Asp Val Phe Arg Glu Ala Gln Arg Ser Tyr Leu  
 95 35 40 45  
 97 Val Phe Leu Thr Ser Gln Phe Cys Tyr Glu Glu Tyr Val Gln Arg Thr  
 98 50 55 60  
 100 Phe Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala  
 101 65 70 75 80  
 103 Ser Val Ala Gly Ala Gly Ala His Ala His Leu Gly Gly Ser Ser Ala  
 104 85 90 95  
 106 Thr Pro Val Gln Gln Ala Gln Ala Ala Ser Ala Gly Thr Gly Ala  
 107 100 105 110  
 109 Leu Ala Ser Ser Ala Pro Ser Thr Ala Val Ala Gln Ser Ala Thr Pro  
 110 115 120 125  
 112 Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala  
 113 130 135 140  
 115 Thr Ala Ala Ala Ser Ala Ala Ala Val Asp Thr Gly Ser Gly Gly  
 116 145 150 155 160  
 118 Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln  
 119 165 170 175

## 122 (2) INFORMATION FOR SEQ ID NO: 3:

## 124 (i) SEQUENCE CHARACTERISTICS:

125 (A) LENGTH: 1038 base pairs

126 (B) TYPE: nucleic acid

127 (C) STRANDEDNESS: double

128 (D) TOPOLOGY: unknown

130 (ii) MOLECULE TYPE: DNA (genomic)

## 132 (vi) ORIGINAL SOURCE:

133 (A) ORGANISM: Epstein-Barr virus

## 136 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

138 ATGCTATCAG GTAACGCAGG AGAAGGAGCA ACAGCCTGCG GAGGTTCCGC CGCCGCGGGC 60  
 140 CAGGACCTCA TCAGCGTCCC CCGCAACACC TTTATGACAC TGCTTCAGAC CAACCTGGAC 120  
 142 AACAAACCGC CGAGGCAGAC CCCGCTACCC TACGCGGCCC CGCTGCCCCC CTTTTCAC 180  
 144 CAGGCAATAG CCACCGCGCC TTCCTACGGT CCTGGGGCCG GAGCGGTCGC CCCGCGCGC 240

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/10/036,729

TIME: 12:02:38

Input Set : N:\Crif3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

```

146 GGCTACTTTA CCTCCCCAGG AGGTTACTAC GCCGGGCCCCG CGGGCGGGGA CCCGGGTGCC 300
148 TTCTTGCCGA TGGACGCTCA CACCTACCAC CCCCACCCAC ACCCCCTCC GGCCTACTTT 360
150 GGCTTGCCGG GCCTCTTTGG CCCCCCTCCA CCCGTGCCTC CTTACTACGG ATCCCACTTG 420
152 CGGGCAGACT ACGTCCCCGC TCCCTCGCGA TCCAACAAGC GGAAAAGAGA CCCCAGAGGAG 480
154 GATGAAGAAG GCGGGGGGCT ATTCCCGGGG GAGGACGCCA CCCTCTACCG CAAGGACATA 540
156 GCGGGCCTCT CCAAGAGTGT GAATGAGTTA CAGCACACGC TACAGGCCCT GCGCCGGGAG 600
158 ACGTGTCCT ACGGCCACAC CGGAGTCGGA TACTGCCCCC AGCAGGGCCC CTGCTACACC 660
160 CACTCGGGGC CTTACGGATT TCAGCCTCAT CAAAGCTACG AAGTGCCCAG ATACGTCCCT 720
162 CATCCGCCCC CACCACCAAC TTCTCACCAG GCAGCTCAGG CGCAGCCTCC ACCCCGGGGC 780
164 ACACAGGCCC CCGAAGCCCA CTGTGTGGCC GAGTCCACGA TCCCTGAGGC GGGAGCAGCC 840
166 GGGAACCTCTG GACCCCGGGA GGACACCAAC CCTCAGCAGC CCACCACCGA GGGCCACCAC 900
168 CGCGGAAAGA AACTGGTGCA GGCCTCTGCG TCCGGAGTGG CTCAGTCTAA GGAGCCCACC 960
170 ACCCCCAAGG CCAAGTCTGT GTCAGCCCAC CTCAAGTCCA TCTTTTGCGA GGAATTGCTG 1020
172 AATAAACGCG TGGCTTGA 1038
174 (2) INFORMATION FOR SEQ ID NO: 4:
176 (i) SEQUENCE CHARACTERISTICS:
177 (A) LENGTH: 345 amino acids
178 (B) TYPE: amino acid
179 (C) STRANDEDNESS: single
180 (D) TOPOLOGY: linear
182 (ii) MOLECULE TYPE: peptide
184 (vi) ORIGINAL SOURCE:
185 (A) ORGANISM: Epstein-Barr virus
188 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
190 Met Leu Ser Gly Asn Ala Gly Glu Gly Ala Thr Ala Cys Gly Gly Ser
191 1 5 10 15
193 Ala Ala Ala Gly Gln Asp Leu Ile Ser Val Pro Arg Asn Thr Phe Met
194 20 25 30
196 Thr Leu Leu Gln Thr Asn Leu Asp Asn Lys Pro Pro Arg Gln Thr Pro
197 35 40 45
199 Leu Pro Tyr Ala Ala Pro Leu Pro Pro Phe Ser His Gln Ala Ile Ala
200 50 55 60
202 Thr Ala Pro Ser Tyr Gly Pro Gly Ala Gly Ala Val Ala Pro Ala Gly
203 65 70 75 80
205 Gly Tyr Phe Thr Ser Pro Gly Gly Tyr Tyr Ala Gly Pro Ala Gly Gly
206 85 90 95
208 Asp Pro Gly Ala Phe Leu Ala Met Asp Ala His Thr Tyr His Pro His
209 100 105 110
211 Pro His Pro Pro Pro Ala Tyr Phe Gly Leu Pro Gly Leu Phe Gly Pro
212 115 120 125
214 Pro Pro Pro Val Pro Pro Tyr Tyr Gly Ser His Leu Arg Ala Asp Tyr
215 130 135 140
217 Val Pro Ala Pro Ser Arg Ser Asn Lys Arg Lys Arg Asp Pro Glu Glu
218 145 150 155 160
220 Asp Glu Glu Gly Gly Gly Leu Phe Pro Gly Glu Asp Ala Thr Leu Tyr
221 165 170 175
223 Arg Lys Asp Ile Ala Gly Leu Ser Lys Ser Val Asn Glu Leu Gln His
224 180 185 190
226 Thr Leu Gln Ala Leu Arg Arg Glu Thr Leu Ser Tyr Gly His Thr Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/036,729

DATE: 02/04/2002

TIME: 12:02:38

Input Set : N:\Crif3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

```

227          195          200          205
229 Val Gly Tyr Cys Pro Gln Gln Gly Pro Cys Tyr Thr His Ser Gly Pro
230          210          215          220
232 Tyr Gly Phe Gln Pro His Gln Ser Tyr Glu Val Pro Arg Tyr Val Pro
233 225          230          235          240
235 His Pro Pro Pro Pro Pro Thr Ser His Gln Ala Ala Gln Ala Gln Pro
236          245          250          255
238 Pro Pro Pro Gly Thr Gln Ala Pro Glu Ala His Cys Val Ala Glu Ser
239          260          265          270
241 Thr Ile Pro Glu Ala Gly Ala Ala Gly Asn Ser Gly Pro Arg Glu Asp
242          275          280          285
244 Thr Asn Pro Gln Gln Pro Thr Thr Glu Gly His His Arg Gly Lys Lys
245          290          295          300
247 Leu Val Gln Ala Ser Ala Ser Gly Val Ala Gln Ser Lys Glu Pro Thr
248 305          310          315          320
250 Thr Pro Lys Ala Lys Ser Val Ser Ala His Leu Lys Ser Ile Phe Cys
251          325          330          335
253 Glu Glu Leu Leu Asn Lys Arg Val Ala
254          340          345
256 (2) INFORMATION FOR SEQ ID NO: 5:
258     (i) SEQUENCE CHARACTERISTICS:
259         (A) LENGTH: 24 amino acids
260         (B) TYPE: amino acid
261         (C) STRANDEDNESS: single
262         (D) TOPOLOGY: linear
264     (ii) MOLECULE TYPE: peptide
266     (vi) ORIGINAL SOURCE:
267         (A) ORGANISM: Epstein-Barr virus
270     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
272 Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro His Asp Thr Ala
273          5          10          15
275 Pro Arg Gly Ala Arg Lys Lys Gln
276          20
279 (2) INFORMATION FOR SEQ ID NO: 6:
281     (i) SEQUENCE CHARACTERISTICS:
282         (A) LENGTH: 30 amino acids
283         (B) TYPE: amino acid
284         (C) STRANDEDNESS: single
285         (D) TOPOLOGY: linear
287     (ii) MOLECULE TYPE: peptide
289     (vi) ORIGINAL SOURCE:
290         (A) ORGANISM: Epstein-Barr virus
293     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
295 Ser Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser Ser Ser Ile
296          5          10          15
298 Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala
299          20          25          30
301 (2) INFORMATION FOR SEQ ID NO: 7:
303     (i) SEQUENCE CHARACTERISTICS:

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/036,729

DATE: 02/04/2002

TIME: 12:02:38

Input Set : N:\Crf3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

```

304      (A) LENGTH: 15 amino acids
305      (B) TYPE: amino acid
306      (C) STRANDEDNESS: single
307      (D) TOPOLOGY: linear
309      (ii) MOLECULE TYPE: peptide
311      (vi) ORIGINAL SOURCE:
312          (A) ORGANISM: Epstein-Barr virus
315      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
317 Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
318      5                               10                               15
320 (2) INFORMATION FOR SEQ ID NO: 8:
322      (i) SEQUENCE CHARACTERISTICS:
323          (A) LENGTH: 15 amino acids
324          (B) TYPE: amino acid
325          (C) STRANDEDNESS: single
326          (D) TOPOLOGY: linear
328      (ii) MOLECULE TYPE: peptide
330      (vi) ORIGINAL SOURCE:
331          (A) ORGANISM: Epstein-Barr virus
334      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
336 Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
337      5                               10                               15
339 (2) INFORMATION FOR SEQ ID NO: 9:
341      (i) SEQUENCE CHARACTERISTICS:
342          (A) LENGTH: 12 amino acids
343          (B) TYPE: amino acid
344          (C) STRANDEDNESS: single
345          (D) TOPOLOGY: linear
347      (ii) MOLECULE TYPE: peptide
349      (vi) ORIGINAL SOURCE:
350          (A) ORGANISM: Epstein-Barr virus
353      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
355 Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser
356      5                               10
358 (2) INFORMATION FOR SEQ ID NO: 10:
360      (i) SEQUENCE CHARACTERISTICS:
361          (A) LENGTH: 12 amino acids
362          (B) TYPE: amino acid
363          (C) STRANDEDNESS: single
364          (D) TOPOLOGY: linear
366      (ii) MOLECULE TYPE: peptide
368      (vi) ORIGINAL SOURCE:
369          (A) ORGANISM: Epstein-Barr virus
372      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
374 Pro Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala
375      5                               10
377 (2) INFORMATION FOR SEQ ID NO: 11:
379      (i) SEQUENCE CHARACTERISTICS:
380          (A) LENGTH: 12 amino acids

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/036,729

DATE: 02/04/2002

TIME: 12:02:39

Input Set : N:\Crf3\RULE60\10036729.raw

Output Set: N:\CRF3\02042002\J036729.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]